HIV whole-genome sequencing now: answering still-open questions

Metzner, Karin J

Abstract: Diversity, evolution, and epidemiology of HIV are directly relevant to HIV transmission and pathogenesis; hence, they play a key role in antiretroviral treatment and vaccine design. Global HIV whole-genome sequencing would provide a treasure chest of data to answer many questions still open in these fields. An article by Berg et al. in this issue of theJournal of Clinical Microbiologydescribes a universal strategy to amplify and sequence heterogeneous HIV whole genomes (M. G. Berg, J. Yamaguchi, E. Alessandri-Gradt, R. W. Tell, J.-C. Plantier, and C. A. Brennan, J Clin Microbiol 54:868-882, 2016,http://dx.doi.org/10.1128/JCM.02479-15).

DOI: https://doi.org/10.1128/JCM.03265-15

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology

COMMENTARY

HIV Whole-Genome Sequencing Now: Answering Still-Open Questions

Karin J. Metzner

Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland; Institute of Medical Virology, University of Zurich, Zurich, Switzerland

Diversity, evolution, and epidemiology of HIV are directly relevant to HIV transmission and pathogenesis; hence, they play a key role in antiretroviral treatment and vaccine design. Global HIV whole-genome sequencing would provide a treasure chest of data to answer many questions still open in these fields. An article by Berg et al. in this issue of the *Journal of Clinical Microbiology* describes a universal strategy to amplify and sequence heterogeneous HIV whole genomes (M. G. Berg, J. Yamaguchi, E. Alessandri-Gradt, R. W. Tell, J.-C. Plantier, and C. A. Brennan, J Clin Microbiol 54:868–882, 2016, http://dx.doi.org/10.1128/JCM.02479-15).

Human immunodeficiency virus (HIV), the cause of pandemic AIDS, is one of—if not the most—extensively investigated viruses. Numerous breakthroughs that have benefits beyond the field of virology were made in HIV research. The most notable, perhaps, is the unprecedented achievement of transforming a deadly infection into a chronic disease by the introduction of combination antiretroviral therapy. But why have previous attempts to develop an effective vaccine been unsuccessful? Because HIV is a master of evolution. As a rapidly replicating RNA virus with a small genome size (<10 kb) and a high mutation rate, all possible single-point mutations could arise in an untreated HIV-infected individual daily. With more than 37 million HIV-infected individuals worldwide today, HIV is one of the world's most genetically diverse pathogens. Currently, two types of HIV exist: HIV-1 and HIV-2. HIV-1 is the cause of the pandemic and is divided into four groups that, on average, feature 37.5% diversity among each other. The main group, M, is further subdivided into nine subtypes and numerous circulating recombinant forms. On average, the percentages of genome diversity among and within HIV-1 subtypes are 14.7% and 8.2%, respectively, whereas much higher diversity can be observed in certain variable regions such as *env* (1). Thus, global molecular surveillance is a necessity to monitor dynamics of HIV diversity and to identify newly emerging HIV strains that may have an impact on diagnostic assay designs and antiretroviral therapy efficacy.

The rapid development of next-generation sequencing (NGS) technology for the past decade could not have come at a better time. Within a few days, billions of sequence reads can be obtained at decreasing costs, and experimental procedures applying NGS technologies are constantly becoming easier. Three years ago, the first pan-HIV-1 strategy to amplify HIV whole genomes using primers in semiconserved regions of all HIV-1 groups and subtypes and sequenced with NGS was published (2). In this issue of *Journal of Clinical Microbiology*, another pan-HIV strategy is described by Berg et al. (3): a more universal approach applying the SMART (switching mechanism at 5′ end of RNA transcript) technology. Here, multiple adaptor-tagged pan-HIV primers are used for reverse transcription of viral RNA into cDNA followed by the addition of nucleotides at the 5′ end of the cDNA. To this extension, another adaptor-tagged primer is hybridized, leading to the switch of template by the reverse transcriptase and the replication of the adaptor. The final cDNA can then be amplified using adaptor-specific primers. This approach captures a higher diversity of genetic variants, albeit with lower sensitivity, meaning higher viral loads are required to successfully obtain HIV whole-genome sequences compared to HIV-specific amplification procedures (3).

In lockstep with the continuing evolution of NGS technologies and experimental procedures, the bioinformatics community has developed many computational tools capable of keeping up with the increasing amount of data and instrumental in their interpretation. However, experimental and computational challenges remain (4), and with the accumulation of sequence data, more sophisticated computer software and infrastructure as well as continuing close collaborations among experimental, clinical, and computational scientists are required.

Nevertheless, NGS technologies enable unprecedented opportunities to investigate HIV diversity. Global molecular surveillance to identify newly emerging HIV strains is only one of numerous possible applications. Large-scale HIV-1 whole-genome sequencing will provide more detailed insights into viral evolutionary processes among and within hosts, infection clusters and transmission chains, fitness landscapes of virus populations, and pathogenesis of different subtypes—to name a few.

New NGS technologies are not only used in science; they are also entering diagnostic laboratories. Benchtop NGS machines are progressively replacing Sanger sequencers since on top of generating consensus sequences, they also have the ability to (ultra-) deep sequence, which is essential in the detection and quantification of low-abundance viruses such as minority drug-resistant HIV variants (5). Although the clinical impact of minority drug-resistant HIV variants remains unclear (6), large-scale and deep HIV genome sequencing could provide the data needed to define, for instance, clinically relevant cutoffs for drug resistance muta-

Accepted manuscript posted online 20 January 2016

Citation Metzner KJ. 2016. HIV whole-genome sequencing now: answering still-open questions. J Clin Microbiol 54:834–835. doi:10.1128/JCM.03265-15.

Editor: Y.-W. Tang

Address correspondence to Karin.Metzner@usz.ch.

tions. This information, however, cannot be inferred from HIV sequences alone. Clinical data from well-characterized HIV-infected individuals are equally important.

A word of caution should be added with respect to the implementation of these new NGS technologies in routine diagnostic settings. Besides a robust laboratory procedure to amplify and sequence a broad range of heterogeneous HIV strains, other requirements have to be fulfilled. Cost-effectiveness, high accuracy and sensitivity, reproducibility, and standardized analysis are crucial considerations for a clinical pipeline to be practical. The cost for routine HIV drug resistance testing might remain the same using Sanger or NGS sequencers, particularly in laboratories with high sample throughputs. Accuracy, sensitivity, and reproducibility in calling consensus sequences and detecting majority drug resistance mutations are noninferior to population sequencing (5, 7, 8). Similarly, standardized analysis and routinely used interpretation algorithms (e.g., the Stanford HIV Drug Resistance Database and geno2pheno) are applicable to Sanger sequencing- or NGS-derived consensus sequences. For these reasons, the application of new NGS methodologies in diagnostic settings should not pose a remarkable challenge as long as the analysis involves consensus sequences only. However, detection and quantification of minority HIV variants are much more difficult and remain complex (4). Numerous experimental procedures and analysis pipelines for different NGS platforms have been developed worldwide, highlighting the efforts undertaken to further explore HIV diversity. As most of these protocols differ in terms of primer design, amplification and sequencing conditions, and data analysis, it is high time to conduct interlaboratory studies in order to consolidate both experience and knowledge (7, 9).

In summary, the application of NGS technologies for HIV whole-genome sequencing will provide unprecedented insights into genomic structures of virus populations that may be directly relevant for clinical and scientific topics, such as personalized antiretroviral treatment and vaccine design. The global monitoring of HIV diversity is important for the surveillance of the pandemic and could have immediate consequences for antiretroviral treatment and diagnostic assay designs. In the long run, these data might even aid in the prediction of virus evolution, such as pathways of HIV to escape immune and drug pressures in HIV-infected individuals. Together with clinical and epidemiological data and host genomic information, HIV whole-genome sequences will allow insights into the highly complex host-pathogen interactions. It is undoubtedly a golden age for research in virus evolution and epidemiology.

ACKNOWLEDGMENTS

I am thankful to Yik Lim Kok, Roger Kouyos, Nicolas Müller, David Seifert, and Osvaldo Zagordi for critical reading of the manuscript and helpful discussions. Support for this article was provided by SystemsX.ch, the Swiss Initiative in Systems Biology (MRD-Project HIV-X).

K.J.M. has received travel grants and honoraria from Gilead, Roche Diagnostics, GlaxoSmithKline, Merck Sharp & Dohme, Bristol-Myers Squibb, ViiV, and Abbott and consulting fees from Gilead and ViiV, and the University of Zurich received research grants from Gilead, Roche, and Merck Sharp & Dohme for studies for which K.J.M. serves as principal investigator.

REFERENCES

1. **Li G, Piampongsant S, Faria NR, Voet A, Pineda-Pena AC, Khouri R, Lemey P, Vandamme AM, Theys K.** 2015. An integrated map of HIV genome-wide variation from a population perspective. Retrovirology **12:** 18. http://dx.doi.org/10.1186/s12977-015-0148-6.
2. **Gall A, Ferns B, Morris C, Watson S, Cotten M, Robinson M, Berry N, Pillay D, Kellam P.** 2012. Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes. J Clin Microbiol **50:**3838–3844. http://dx.doi.org/10.1128/JCM.01516-12.
3. **Berg MG, Yamaguchi J, Alessandri-Gradt E, Tell RW, Plantier J-C, Brennan CA.** 2015. A pan-HIV strategy for complete genome sequencing. J Clin Microbiol **54:**868–882. http://dx.doi.org/10.1128/JCM.02479-15.
4. **Beerenwinkel N, Gunthard HF, Roth V, Metzner KJ.** 2012. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. Front Microbiol **3:**329. http://dx.doi.org/10.3389/fmicb.2012.00329.
5. **Lapointe HR, Dong W, Lee GQ, Bangsberg DR, Martin JN, Mocello AR, Boum Y, Karakas A, Kirkby D, Poon AF, Harrigan PR, Brumme CJ.** 2015. HIV drug resistance testing by high-multiplex "wide" sequencing on the MiSeq instrument. Antimicrob Agents Chemother **59:**6824–6833. http://dx.doi.org/10.1128/AAC.01490-15.
6. **Li JZ, Kuritzkes DR.** 2013. Clinical implications of HIV-1 minority variants. Clin Infect Dis **56:**1667–1674. http://dx.doi.org/10.1093/cid/cit125.
7. **Simen BB, Braverman MS, Abbate I, Aerssens J, Bidet Y, Bouchez O, Gabriel C, Izopet J, Kessler HH, Stelzl E, Di Giallonardo F, Schlapbach R, Radonic A, Paredes R, Recordon-Pinson P, Sakwa J, St John EP, Schmitz-Agheguian GG, Metzner KJ, Daumer MP, 454 HIV Alpha Study Group.** 2014. An international multicenter study on HIV-1 drug resistance testing by 454 ultra-deep pyrosequencing. J Virol Methods **204:**31–37. http://dx.doi.org/10.1016/j.jviromet.2014.04.007.
8. **Stelzl E, Proll J, Bizon B, Niklas N, Danzer M, Hackl C, Stabentheiner S, Gabriel C, Kessler HH.** 2011. Human immunodeficiency virus type 1 drug resistance testing: evaluation of a new ultra-deep sequencing-based protocol and comparison with the TRUGENE HIV-1 genotyping kit. J Virol Methods **178:**94–97. http://dx.doi.org/10.1016/j.jviromet.2011.08.020.
9. **St John EP, Simen BB, Turenchalk GS, Braverman MS, Abbate I, Aerssens J, Bouchez O, Gabriel C, Izopet J, Meixenberger K, Di Giallonardo F, Schlapbach R, Paredes R, Sakwa J, Schmitz-Agheguian GG, Thielen A, Victor M, Metzner KJ, Daumer MP, HIV Alpha Study Group.** 2016. A follow-up of the multicenter collaborative study on HIV-1 drug resistance and tropism testing using 454 ultra-deep pyrosequencing. PLoS One **11:**e0146687. http://dx.doi.org/10.1371/journal.pone.0146687.